The Thesis Committee for Alexandra Estrella
Certifies that this is the approved version of the following thesis:

Informational Masking of Multi Talker Babble in English Vowel Identification for Spanish-English Bilinguals

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor: ______________________________
Chang Liu

____
Bharath Chandrasekaran

Informational Masking of Multi Talker Babble in English Vowel Identification for Spanish-English Bilinguals

by

Alexandra Estrella, B.S., B.A.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2016

Dedication

I'd like to dedicate this thesis and degree to my family. Thank you for all the help and support, including all the advice, jokes, and of course, the academic help that helped me get to this point!

Acknowledgements

Thank you to everyone in my cohort—even if we only exchanged a few words, every single one of you shaped my graduate school experience in a way. Special thanks to Liz and Louisa for being great support in the last few weeks of this arduous trek!

Also, thank you Jingjing for being available to help me at the lab when I couldn’t read my data analysis or when I needed a pair of expert eyes to read what I had written! I really appreciated the help.

Abstract

Informational Masking of Multi Talker Babble in English Vowel Identification for Spanish-English Bilinguals

Alexandra Estrella, M.A.

The University of Texas at Austin, 2016

Supervisor: Chang Liu

Speech perception studies with bilinguals have demonstrated that bilinguals perform comparably to native speakers in listening conditions during quiet conditions. However, when the listening conditions included different types of noise, and different SNRs, bilinguals are seen to have difficulties and perform lower than native speakers when tested in their L2. With Spanish-English bilinguals becoming a large part of the U.S. population, the present study investigated their speech perception abilities using English vowels in different quiet and noise conditions. The participants were controlled for their age of acquisition of English in order to determine if the amount of exposure to the language affected their overall performance. In addition, the amount of informational masking was evaluated using comparisons with the babble and temporally modulated noise conditions. Results indicated that the later bilinguals experienced more difficulties throughout the different conditions when compared to the simultaneous and early bilinguals, but significance levels were only noted for a few of the conditions. Additionally, there were no major effects for informational masking.

Table of Contents

List of Tables

List of Figures

INTRODUCTION AND LITERATURE REVIEW

There are nearly 60 million individuals in the United States who currently speak a language other than English at home (Callahan & Gándara, 2014). Out of these 60 million individuals, about 62% of them speak Spanish (Callahan & Gándara, 2014). Between 1980 and 2007 there was a 210.8% increase in the number of individuals who speak Spanish at home, and this type of growth is expected to increase in the following years (Shi, 2014).

With an increasing Spanish and English bilingual population, several studies have looked at the speech and language development of bilinguals, and the overall benefits of bilingualism. Bilingual speakers have fewer difficulties when learning an additional language close to their L1 or L2, and they benefit from increased job opportunities and bonuses (Callahan & Gándara, 2014). However, there is a need to investigate the effects of bilingualism on the speech perception of English, which in most cases is the L2. Investigating speech perception of English is important because while there is a large number of bilinguals in the United States, the mainstream language is still English and it comprises the majority of what individuals are exposed to outside of their home environment, at school, and in the workplace (Shi, 2014). Although there are several advantages to being bilingual, it is still unknown whether these benefits translate to speech perception abilities. Therefore, it is important to consider the implications that difficulties perceiving speech stimuli in noise has for bilingual individuals in the school, work, or everyday setting. These different environments have different types of noise that can be emulated in the research setting, therefore, the impact of noise backgrounds on English speech perception for monolingual and bilingual speakers have been assessed by researchers.

However, investigating the specific speech perception abilities of Spanish-English bilinguals is important because of the exponential growth in Spanish-English bilinguals. These individuals are a growing population whose speech perception needs to be assessed with appropriate measures and norms that take into account the difficulties in noise conditions.

SPEECH PERCEPTION IN BILINGUALS

A bilingual individual initially interprets sounds in a targeted language using their L1 knowledge and sound system, making the similarities or differences between both languages an important factor in speech perception (Garcia Lecumberri, Cooke, & Cutler, 2010; Bradlow, Clopper, Smiljanic, & Walter, 2010). Despite this difference to monolinguals in language processing, bilingual individuals have been found to have the same listening abilities as native speakers in quiet conditions (Mayo et al. 1997; Guan et al. 2015; Tabri, About Chacra, Pring, 2010; Vineetha, Suma, Nair, 2013). When there is background noise in the environment, both native and non-native individuals experience challenges when recognizing the targeted speech. However, stimuli presented with competing noise is more difficult for bilinguals to perceive accurately when compared to monolinguals (Mi et al., 2013; Vineetha, Suma, Nair, 2013). Native listeners do not experience the same difficulties as bilinguals do due to the "native advantage" (Jin & Liu, 2014). This is the ability of native listeners to understand speech stimuli when non-natives experience difficulties due to the listening conditions (Jin & Liu, 2014). Native speakers use high-level information, such as suprasegmentals, more effectively to compensate for their problems perceiving a signal (Broersma & Scharenborg, 2010). Individuals who are non-natives to a language and who are bilingual have demonstrated difficulties when presented with different speech stimuli in their non-native language under a variety of noise conditions.

Studies have specifically looked at the performance of bilinguals with existing hearing tests and different informal conditions. Weiss and Dempsey (2008) assessed Spanish-English bilinguals using the English and Spanish versions of the Hearing in Noise Test (HINT). Their study contained four test conditions: in quiet, in noise, and in noise presented on 90 degrees to the left and to the right of the participant. The participants of the study were labeled as Early Bilinguals (EB) or Late Bilinguals (LB). They each had to repeat the HINT sentences presented to them. The results indicated that all participants performed better in the Spanish HINT with Spanish being their L1. The LB group performed better on the Spanish HINT but the EB group performed better on the English HINT. When the data was analyzed using bilingual age as a factor, e.g., the bilingual age increased, performance on the Spanish HINT was reduced. No significant correlations were found between bilingual age and the English HINT.

Shi and Sanchez (2010) measured recognition of English and Spanish words in Spanish-English bilinguals who reported their native language was Spanish. These individuals acquired English at birth or up until their 30s, and were not all balanced bilinguals. The words were presented in quiet and in 0 and +6 dB SNR. Age of acquisition had an effect on their performance on the English test, their L2, but not in the Spanish test, their L1. These findings are consistent with Mayo et al. (1997)'s finding that length of foreign language study, the age of the listener, and the listening conditions significantly affect the speech perception abilities of bilingual individuals.

Overall, according to Shi (2014), Spanish-English Bilinguals perform comparably to English monolinguals when it comes to speech recognition tasks of phonemes, words, or sentences in quiet. However, when there is competing noise, particularly at low SNRs, the speech perception of bilingual listeners is affected more negatively than that of monolingual listeners. Presently, there are not enough studies that specifically investigate the effects of different combinations of noise and SNR in the speech perception in Spanish-English bilinguals. Additionally, studies have primarily compared the performance of bilinguals to monolinguals, and have not thoroughly investigated the differences in performance between a range of bilinguals. Several factors influence the listening abilities of bilinguals in addition to the stimuli and listening conditions and these factors are described as follows.

AGE OF FIRST EXPOSURE

One common term used to describe when an individual begins to learn a second language is age of acquisition (AoA). AoA refers to the age at which an individual began to learn a language or began to be immersed in a language (Birdsong, 2006). However, when it comes to simultaneous bilinguals and individuals who begin learning a language in school, Age of first Exposure (AoE) is most commonly used (Birdsong, 2006). In individuals who grow up in Spanish-speaking households, this refers to the age at which they began to learn English. For many individuals this may be the time when they begin school and are placed in mainstream English classrooms or bilingual classrooms. For others, their first exposure to English correlates with their AoA, or when they immigrated to the United States and received exposure to English by living in a new country (Birdsong, 2006).

AoA and AoE are important factors in determining overall competency in a language. According to the Critical Period Hypothesis (CPH) individuals who learn a language before puberty are more likely to master the language with native-like fluency (Birdsong, 1999). This is not to say that individuals who learn a language after puberty will not reach native-like levels of fluency, but research has pointed that as AoA increases, individuals are more likely to make grammatical and morphosyntactic errors that decrease their overall proficiency (Birdsong, 2006). The CPH does not specifically address the effects on speech perception. However, the relationship between the production and perception of an L2 are considered to be complex but related: later exposure to an L2 will hinder speech perception in that language (Hisagi, Garrido-Nag, Datta, Shafer, 2015). On the other hand, simultaneous and early bilinguals were found to be able to adapt and shift their perception to fit the target language (Hisagi, Garrido-Nag, Datta, Shafer, 2015). If AoA and AoE play a part in the overall proficiency of an individual, one research question is whether it also plays a role in her/his speech perception.

Tabri et al. (2011) reported that although bilinguals who acquired English at an early age do not perform as well as monolinguals, they perform better than bilinguals who learned the language at a later age in noise conditions. According to Bovo and Callegari (2009), if a bilingual has not completed their development of their L2, it is difficult for them to understand speech stimuli in that language. The children in their study who had "low linguistic competence" and who had only been studying the target language for 1-3 years were at a greater disadvantage when speech was presented in noise with energetic masking (Bovo & Callegari, 2009). Therefore, lower proficiently in a language may result in more difficulties with speech perception in noise conditions.

SPEECH STIMULI

The type of speech stimuli also affects the difficulty for bilinguals to perceive the target accurately. Initially, individuals need to transform acoustic signals into phonemes (Wilson & Iacoboni, 2006). Then, listeners rely on the characteristics of speech sounds, including the spectral and temporal characteristics in addition to syntactic structure, semantic context, and pragmatic factors when available (Morrill, Baese-Berk, Heffner, & Dilley, 2015). Single phoneme stimuli are brief and do not have those additional cues, making them harder to perceive accurately, compared to words and sentences in which listeners are able to repair speech signals using additional syntactic and semantic cues (Morrill, Baese-Berk, Heffner, & Dilley, 2015). That is, single phonemes can be harder to perceive because they are stand-alone and do not carry semantic cues that words and sentences contain (Broersma & Scharenborg, 2010).

LISTENING CONDITIONS

In addition to the speech stimuli presented, the listening conditions affect how easily speech can be perceived. The listening conditions include the type of noise being presented and the variations in signal-to-noise (SNR). Several studies have investigated the effects of noise and determined that the competing noise will affect speech perception in bilinguals (Broersma & Scharenborg, 2010). Although different types of noise and SNRs affect all listeners, they will have a more detrimental effect on the overall perception abilities of the bilingual (Broersma & Scharenborg, 2010).

Type of Noise

The competing noise presented during a speech perception task affects the difficulty to understand the stimuli (Brouwer, Van Engen, Calandruccio, Bradlow, 2011). There are different types of noise, including multi-talker babble (MTB) and long-term speech-shaped noise (LTSS). Mi et al. (2013) investigated the vowel identification abilities in English and Chinese listeners under different types of noises. There was one English Native (EN) group while the Chinese-native listeners were from the U.S. (CNU) or China (CNC). Twelve different English vowels were presented in long-term speech-shaped noise (LTSSN) and in muti-talker babble (MTB). The EN group outperformed both the CNU and CNC groups, but both Chinese native groups had very little difference in their identification performance (Mi et al., 2013). When the vowels were presented in the LSSTN and MTB conditions, performance decreased for all groups but the EN group still performed better than the Chinese native groups.

MTB is a masker that listeners are most likely to experience in their everyday life. Multi-talker babble is also variable, since there can be a different number of speakers, the speakers may be different genders, and the babble can be in different languages (Silbert, de Jong, Regier, Albin, & Hao, 2014). The effects of babble on speech perception are important because when the stimuli and maskers are similar there is more confusion and difficulty (Rhebergen, Versfeld, Dreschler, 2005). There are several factors related to MTB that vary the difficulty for perceiving speech stimuli. First, MTB is acoustically similar and related to the speech signals being presented. This makes it harder for the listener to pay attention to the signal and not the masker (Cooke et al. 2008). As the number of speakers in the MTB increases up to eight in general, the speech stimuli become harder to perceive (Marchegiani & Fafoutis, 2015; Snell et al. 2002). Another factor that will affect how well the speech stimuli is perceived is familiarity with the language of the masker. For example, Dutch listeners had a harder time perceiving speech stimuli when the MTB was in Dutch, but not as much when the MTB was in Swedish (Cooke et al. 2008). Spanish-English bilinguals were found to have similar difficulties with competing MTB in both English and Spanish, even when English was their L2 (Cooke et al. 2008).

LTSS noise is another commonly used noise that is spectrally matched with multi-talker babble (Liu & Eddins, 2008). In the present study, the LTSS noise is babble modulated and is acoustically similar to the MTB. Modulated noise, such as the LTSS used in the present study, has been found to be a less effective masker than MTB due to the lack of informational masking (Simpson & Cooke, 2005). However, LTSS noise also allows for a comparison of the amount of difficulties caused by the acoustic characteristics of MTB and to calculate the amount of informational masking.

Informational Masking

In broad terms, masking occurs when the target stimuli are degraded and harder to discriminate due to a competing signal (Watson, 2005). The different types of noise will result in different masking effects: energetic and informational. In energetic masking there are portions of the speech stimuli that are inaudible due to the masker (Cooke, Garcia Lecumberri, & Barker, 2007). Informational masking (IM) includes misallocation of audible masker components, competing attention of the masker, a higher cognitive load, or inference from the masker (Cooke, Garcia Lecumberri, & Barker, 2007). Informational masking occurs when the signal and the masker are audible, but they are hardly distinguishable, and can be calculated by comparing performance in MTB and LTSS noise (Brungart, 2000).

Signal-to-Noise Ratio (SNR)

SNR refers to the difference between the signal being presented and the competing noise in the environment (Gillam, Marquardt, & Martin, 2011). A positive SNR would indicate that the target signal being presented is of a higher intensity than the noise, making it easier to discern the signal. A negative SNR would indicate that the target signal is of a lower intensity than the noise. If the SNR is negative, the signal would be difficult to discern. Ideal environments would have a positive SNR, while negative SNRs create environments where speech perception becomes difficult.

Performance on speech perception tasks increase as the SNR increases, but Brungart (2000) found that this effect is seen in energetic masking but not informational masking. SNR begins to have a negative effect on the EM of speech perception when SNR drops from –3 dB to –12 dB (Brungart, 2000). On the other hand, informational masking does not always increase when SNRs decrease from 0 to –12 dB (Brungart, 2000). Less IM at a lower SNRs affect all listeners, including early bilinguals (Shi & Sanchez, 2010).

HYPOTHESIS

Research has looked at the overall speech perception abilities of bilingual individuals. Bilingual individuals perform similarly to native listeners in quiet conditions (Mayo et al. 1997; Guan et al. 2015; Tabri, About Chacra, Pring, 2010). Furthermore, in conditions with noise, monolinguals outperform bilingual listeners. However, studies have not focused on the performance of English vowel identification for Spanish-English bilinguals with different Ao of English.

The current study would add to the body of literature of speech perception in Spanish-English bilinguals by investigating their performance of English vowel perception in different types of noise and SNRs. Based on previous research, it is expected that Spanish-English bilinguals will experience difficulties when listening to target stimuli in several noise conditions and increasingly negative SNR conditions. It is expected that the individuals who were exposed to or acquired English at a later age will face more difficulties and achieve lower perception scores than the individuals who learned English earlier. Even if the late bilinguals have had fewer years of exposure to

English, they are also likely to have less exposure to noisy conditions in English, so they are expected to perform worse in all conditions. Additionally, the six-talker babble and corresponding LTSS noise should present the participants with more difficulties than the two-talker babble and corresponding LTSS conditions due to increased number of talkers in babble.

METHODS

PROJECT DESIGN

The project was constructed and run using SykofizX software. SykofizX is a software program developed by Tucker-Davis Technologies. The program is used to run psychophysical experiments in audition, and was specifically designed for speech perception and psychoacoustic studies (SykofizX Manual). The current project was designed to investigate vowel identification of English-Spanish bilinguals when the sounds were presented in different noise conditions. The SykofizX program facilitated the changes in types of noise, including babble or temporally modulated noise, and different SNRs.

The project was primarily separated into two separate sets, the quiet and noise conditions. In total, there were 10 conditions that each participant completed. The first set were two conditions in quiet. The second set were eight conditions with different noise, either multi-talker babble (MTB) or temporally modulated long-term speech shaped (LSST-TM) noise. The noise conditions included two- and six-talker babbles with corresponding LTSS noises. The SNR was set at -3 and -9 dB. There were a total of eight noise conditions (4 types of noises x 2 SNRs).

The participants completed two quiet conditions at separate times, one at the beginning and the other at the end. Their performance in the first quiet condition was compared to their performance in the second quiet condition to determine if there were a learning effect that would influence their performance throughout the experiment. The eight noise conditions were randomized between the two quiet conditions.

During each condition, the participant would hear one of 12 different vowels /æ, ɛ, e, i, ɪ, ɑ, ɔ, o, ʌ, u, ʊ, ɜ/ and was asked to click on the corresponding button with the correct word on the SykofizX program. Each vowel was presented within a word, e.g., had, hayed, hawed, heed, head, hid, heard, hod, hood, whod, hoed, hud. Participants were given the opportunity to complete up to three different training sessions so they could become comfortable with the experiment set up.

PARTICIPANTS

A total of 33 individuals participated in the experiment. The participants were 14 males and 17 females with an average age of 21.93 years. Every participant had to be a Spanish-English bilingual whose first language was Spanish and who either simultaneously learned English or who learned English later throughout their life. Each participant reported normal hearing and did not indicate knowing a third language.

Each participant was paid for their time completing the experiment. All individuals were students enrolled in courses at the University of Texas at Austin or were employees at the university.

Table 1. Participant Breakdown

	Male	Female	Average Age
Simultaneous	6	5	22.72
Early	4	7	20.72
Late	4	7	22.36
Total	14	17	21.93

Data Collection

Once each participant was seated in the booth, they would receive instructions on how to work on the SykofizX program. Each participant used Sony headphones and heard the sound stimuli through the right headphone. Each participant started the experiment by completing three training sessions. Each participant was instructed to answer as fast and as best as they could after the stimuli were presented. Participants were encouraged to take breaks between conditions.

Data Analysis

The data for each participant was exported from SykofizX as a TXT file. The data from the TXT file was pasted into an Excel file template that was set up to calculate the percentage accuracy for each vowel and overall accuracy for each of the 10 conditions. Then, the data percentages for each participant were loaded into a single Excel file. The data for each participant was labeled with their code name and with the group each participant belonged to. There were three groups in the experiment, simultaneous, early, and late bilinguals.

The participants were separated into groups depending on their AoA of English. Group 1 (G1) were simultaneous bilinguals who were defined as the participants who learned English and Spanish from birth up to three years of age before they started school. Group 2 (G2) were early bilinguals were defined as the participants who learned Spanish first and later acquired English from age four up to age six. Finally, Group 3 (G3) were late bilinguals who were defined as participants who learned Spanish first and acquired English from age seven and up. The oldest age when a participant began to learn English was reported to be 18 years.

For statistical purposes, the percentage data was converted to rationalized arcsin units (RAU; Studebaker, 1985). The noise and quiet data were analyzed using Statistica software. A one-way ANOVA was run for the quiet data. A Tukey Post-Hoc analysis was also run to examine the significance between the performance of each group. The noise data was analyzed using multi-factor ANOVAs that examined the overall effects of the listener group, the type of noise, and SNR. Additional Tukey Post-Hoc analysis were also run when necessary.

The data for each participant was also analyzed to determine the effects of informational masking. The amount of informational masking was obtained by subtracting the percentages for the LTSS-TM noise from the MTB percentages. This data was also examined with multi-factor ANOVAs to determine the effects of listener group, noise type, and SNR. Tukey Post-Hoc analyses were also run were appropriate.

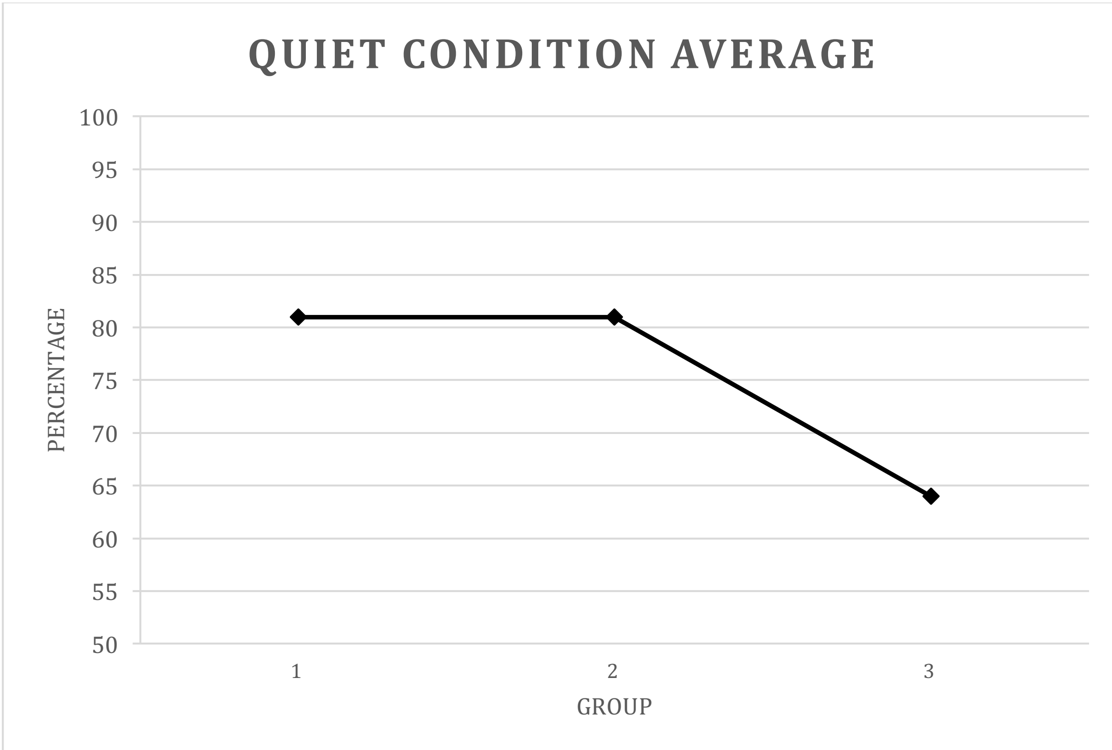
RESULTS

The primary purpose of this study was to determine the effects of age of acquisition and exposure to English on English vowel perception in noise. The participants were separated into three different groups based on their AoA of English.

QUIET CONDITIONS

Averages for both quiet conditions were calculated for each group and plotted in Figure 1. The participants were grouped based on their Ao of English. G1 is the simultaneous bilinguals who learned both English and Spanish since birth. G2 is the early bilingual group with individuals who learned English from age 3 to age 6. G3 is the late bilingual group with individuals who learned English from age 7 and up. G1 and G2 performed equally in their overall average for both quiet conditions. G3 performed lower than both groups, at 64% accuracy for both quiet conditions. For the first quiet condition, G2 scored the highest at 82%, then G1 with 80%, and lastly G3 with 61%. In the second quiet condition, G1 scored the highest with 82%, then G2 with 81%, and lastly G3 with 66%. The differences between conditions one and two were between one and five percent. Figure 1 shows a linear graph with the averages for the three groups.

Figure 1. Quiet Conditions



A one-way ANOVA was run to determine if there was any significance between the performance of each group. Table 2 shows that there was a significant difference in the performance of listener groups, (p = 0.0042). An additional analysis was run to determine where the significant differences were found.

Table 2. Univariate Test for Quiet Conditions

	Sum of Squares	df	Mean Square	F	Sig.
Group	2750.0	2	1375.0	6.5876	0.004249

Note: df: Degrees of Freedom, F: F-value, Sig.: Significance

Results of the Tukey Post-Hoc test determined which groups were significantly different from each other. There was no significant difference between G1 and G2 (p = 0.99). However, G3 was significantly lower than G1 and G2 with a significance noted for both groups ($p = 0.109$, $p = 0.009$).

Noise Conditions

The noise conditions were analyzed initially as the overall averages for each group, for each type of noise and SNR. The results for each type of noise are graphed in Figures 2 and 3. Figure 2 shows the performance of each group for an SNR of -3 dB. Figure 3 shows the performance of the three groups for the SNR of -9 dB.

Figure 2. Noise Conditions -3 dB SNR

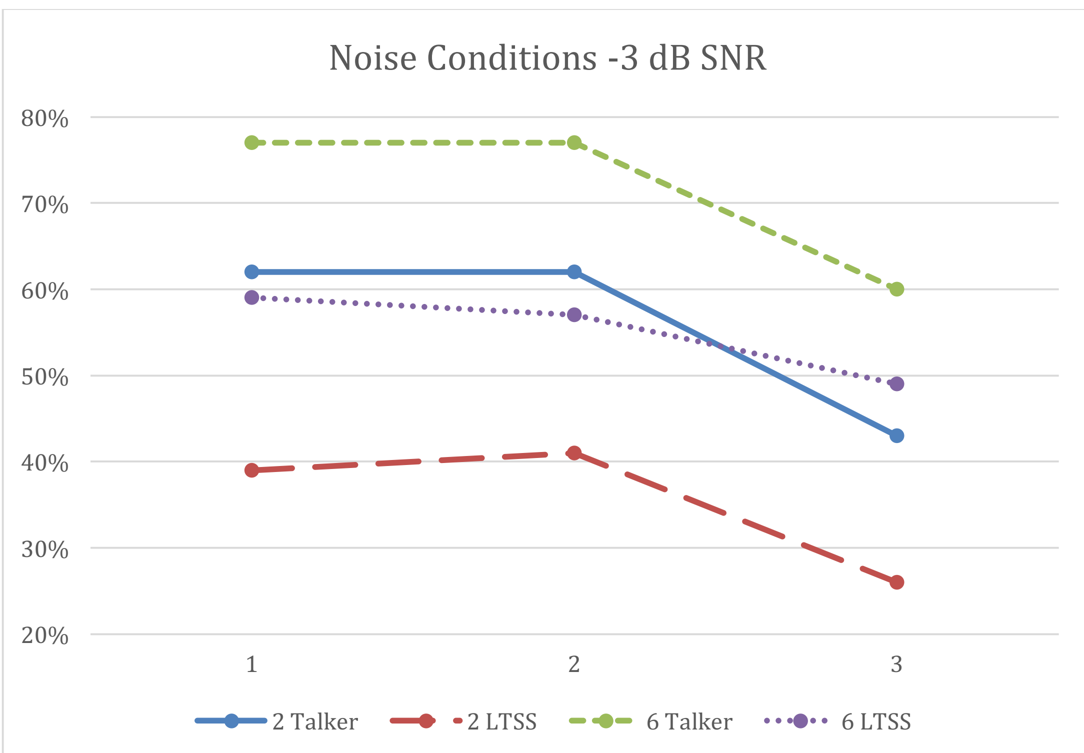
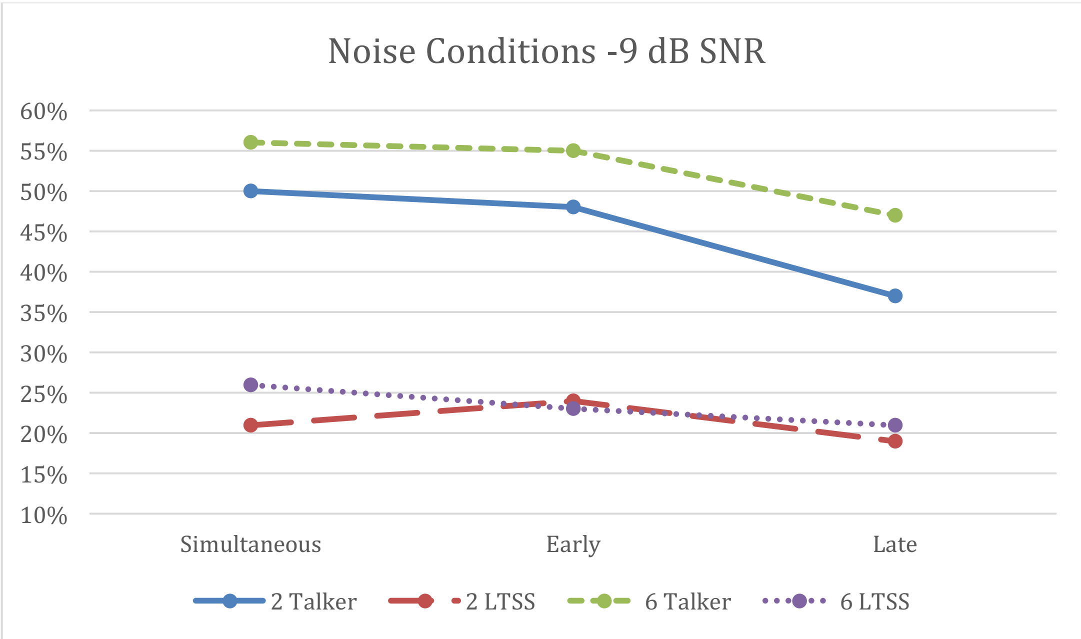


Figure 3. Noise Conditions -9 dB SNR



The data of the noise conditions were analyzed using a three-way ANOVA (Table 3). There was a significance based on the group, ($p = 0.04$). There was also a significance based on the SNR during the conditions, ($p = 0.00$). A third significance was noted based on the Type of Noise, ($p = 0.00$).

Table 3. Three-Way ANOVA for all Noise Conditions

	Sum of Squares	df	Mean Square	F	Sig.
Group	128153	2	64076	3.5912	0.039971*
SNR	761634	1	761634	197.2669	0.000000*
SNR*Group	631	2	316	0.0818	0.921692
Type of Noise	61939	3	20646	14.1875	0.000000*
Type of Noise*Group	9208	6	1535	1.0545	0.395894

Note: df: Degrees of Freedom, F: F-value, Sig.: Significance
*denotes a statistical difference

A Tukey Post-Hoc Analysis was run to compare the performance among the three groups in all noise conditions. The results show that there was no significant difference between each groups performance, although G3 neared a significant difference with G1, ($p = 0.056$) and with G2 ($p = 0.084$).

A Two-Way ANOVA (Table 4) was run for each individual noise condition to determine if there was any significance between the groups or SNR. There was a significance for all of the noise conditions and SNR, ($p = 0.00$), so additional analyses were run to determine if there was a significance between the different noise and SNR combinations and the three different groups.

The additional analyses were run using Two-Way ANOVAs and by looking at the different SNRs (Table 5). There were significant differences for a -3 dB SNR in Two-Talker Babble, Six-Talker Babble, and 2 LTSS noise, (p = 0.004, 0.049, 0.004). However, when doing a detailed analysis of significance between SNR and groups, the value was not found to be significant as the p-value needs to be ($p = >.001$) in order to be significant.

Additionally, there was a significance, ($p = 0.028$), in the 2 LTSS noise condition and the different groups. A Tukey Post-Hoc Analysis was run to determine the difference between the groups (Table 7).

Table 4. Two-Way ANOVA for Each Noise Condition

		Sum of Squares	df	Mean Square	F	Sig.
Two-Talker Babble	Group	20625.1	2	10312.6	2.3100	0.116657
	SNR	48816.1	1	48816.1	59.2498	0.000000*
Six-Talker Babble	Group	20625.1	2	10312.6	2.3100	0.116657
	SNR	48816.1	1	48816.1	59.2498	0.000000*
2 LTSS Noise	Group	47143	2	23571	4.0281	0.028212*
	SNR	125776	1	125776	424.6205	0.000000*
6 LTSS Noise	Group	16376	2	8188	2.3097	0.116692
	SNR	292093	1	292093	302.2187	0.000000*

Note: df: Degrees of Freedom, F: F-value, Sig.: Significance
*denotes a statistical difference

Table 5. Two-Way ANOVAs for Noise Conditions by SNR

	SNR	Sum of Squares	df	Mean Square	F	Sig.
Two-Talker Babble	-3	2750.0	2	1375.0	6.5876	0.004249*
	-9	17941.7	2	8970.8	2.7630	0.079184
Six-Talker Babble	-3	24819.5	2	12409.7	3.3398	0.049029*
	-9	2262.3	2	1131.1	0.71933	0.495290
2 LTSS Noise	-3	45645	2	22822	6.4903	0.004547*
	-9	9209	2	4605	1.7498	0.191090
6 LTSS Noise	-3	14118	2	7059	2.2165	0.126526
	-9	4237.6	2	2118.8	1.59659	0.219325

Note: df: Degrees of Freedom, F: F-value, Sig.: Significance
*denotes a statistical difference

A Tukey Post-Hoc Analysis was run in order to determine what the amount of significance between the groups for the Two-Talker babble (Table 6). The results demonstrated a significant difference between G1 and G3 ($p = 0.049$). A Tukey Post-Hoc Analysis was also run to determine what the significance between the groups for the two-talker LTSS noise (Table 7). There was a significant difference between G1 and G3 ($p = 0.043$). The differences for Six-Talker Babble and the corresponding LTSS noise did not result in any significant difference between the groups.

Table 6. Tukey Post-Hoc Test for Two-Talker Babble

	p-value		
	Simultaneous	Early	Late
Simultaneous		0.970478	0.049635*
Early	0.970478		0.081400
Late	0.049635*	0.081400	

Note: * denotes statistical difference

Table 7. Tukey Post-Hoc Test for 2LTSS Noise

	p-value		
	Simultaneous	Early	Late
Simultaneous		0.985418	0.042895*
Early	0.985418		0.061241
Late	0.042895*	0.061241	

Note: * denotes statistical difference

INFORMATIONAL MASKING

Analyses were run to determine if the effects of informational masking were significant between groups. A Two-Way ANOVA was run to analyze the effects of informational masking between the three groups and SNR. No significant difference was noted among the three groups. Results indicated a significant difference for SNRs, as shown in Figures 4 and 5.

Figure 4. Informational Masking for 2 Talker

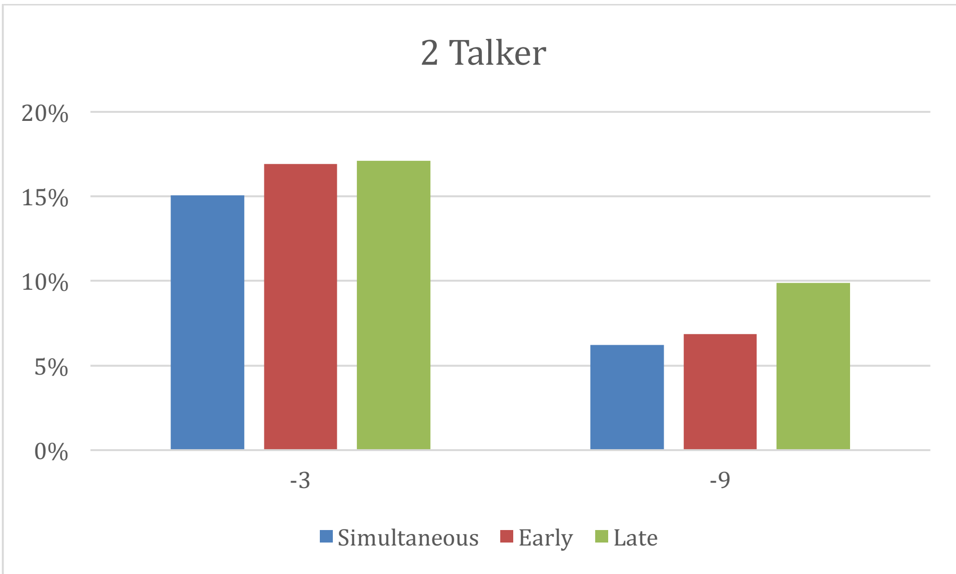
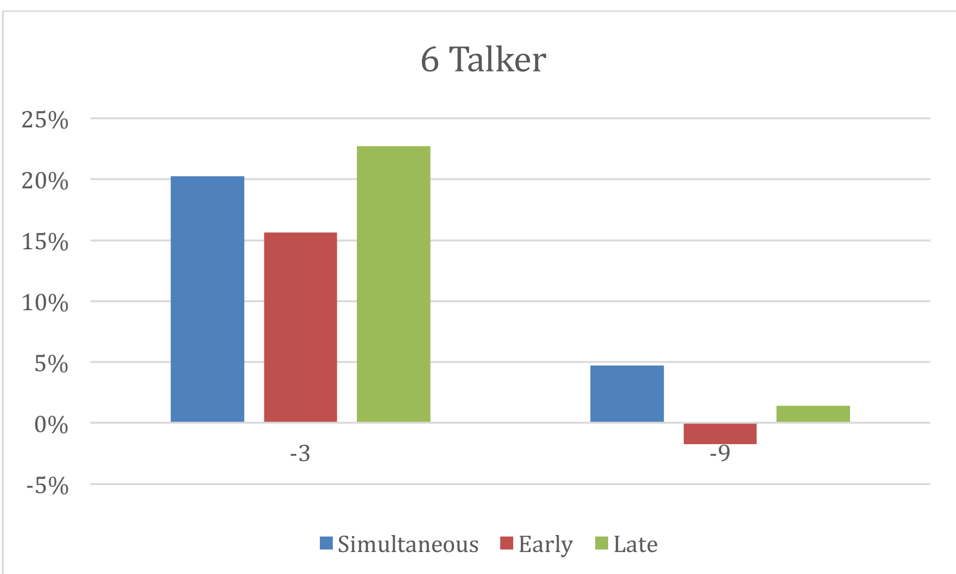


Figure 5. Informational Masking for 6 Talker



DISCUSSION

This study measured the speech perception abilities of Spanish-English bilinguals when they were presented with English vowels in quiet and noise conditions. The current hypothesis was that bilinguals who acquired English later in age would have greater difficulties and would not perform as well as the bilinguals who learned English simultaneously and earlier in life.

All of the participants were assigned to a group depending on their age of English language learning. For the majority of the participants, this was at birth or when they began school. A few participants were exposed to and began learning English when they were older due to migration or delayed language immersion. The participants' performance on each of the conditions was calculated through percentages that were later converted to RAU measurements. The statistical analyses showed that the late bilingual group who learned English between age 7 and 18, G3, performed lower than the simultaneous and early bilingual groups who learned English since birth or between age 4 to 6, G1 and G2. These findings were consistent with the hypothesis.

LATE BILINGUAL PERFORMANCE

One primary hypothesis was that the late bilinguals would perform lower in the L2 speech perception tasks due to their language influence and decreased experience in adverse listening conditions. Previous research has demonstrated that bilinguals have more difficulties with speech perception tasks compared to monolinguals, but also that later bilinguals perform lower compared to earlier bilinguals. In the current study, G3 was made up of the late bilinguals who learned English at age seven and later, having lower scores than G1 and G2, early bilinguals.

In the quiet conditions, the simultaneous and early bilingual groups (G1 and G2) performed similarly. However, the late bilingual group (G3) performed significantly lower than G1 and G2. These differences in the quiet conditions suggest that the late bilinguals had more difficulties than the earlier bilinguals. This can be attributed to the differences in their age of acquisition and exposure to English.

In the noise conditions, there were significant differences between G1 and G3 in the two-talker babble and corresponding LTSS noise. These differences between the simultaneous bilingual group (G1) and the late bilingual group (G3) indicate that the age of acquisition or exposure to a language can also play a factor in speech perception in noise conditions.

Interestingly, the early bilingual group (G1) significantly outperformed the late bilingual group (G3) for the two-talker babble, while there was no significant group difference for the six-talker babble. MTB with more speakers tends to increase the level of difficulty (e.g., greater IM) when perceiving speech stimuli, but this was not noted with the current participants. Their performance in the two-talker babble were lower than in the six-talker babble. Lastly, although G1 and G2 performed similarly in several conditions, the differences between G2 and G3 were not significant.

INFORMATIONAL MASKING

The effects of informational masking on the overall performance of the different groups was also investigated in this study. The results demonstrated through the IM of the three groups that they did not perform significantly differently from each other despite the different vowel identification scores in quiet and noise across the three groups.

Although there were no significant differences in the performance of each group, or between the different types of noise, the groups had higher overall percentages in the LTSS noise conditions than they did in the MTB conditions. These differences in performance can be attributed to the fact that the MTB is more similar to the speech signals that were presented, leading to greater IM.

LIMITATIONS

One limitation of this study is the distributions for the groups used for the analysis. The first group consisted of simultaneous bilinguals, or individuals who reported learning both Spanish and English from birth. The second group consisted of early bilinguals, or individuals who reported that they learned English between three to six years of age when they began school. The third group consisted of late bilinguals, but it covered a larger span, starting at seven years and up until 18 years of age.

An additional limitation of this study was the lack of a control group made up of monolingual English speakers. A control group would have been helpful for drawing comparisons of the performance between monolinguals and bilinguals, especially for comparing how a sample of college individuals would have performed differently from each other.

Additional Considerations For Future Research

Additional research investigating the speech perception of bilinguals should have a larger sample. Along with a larger sample, it would be beneficial to have a "Middle Bilingual" group that breaks up the current late bilingual group. The current late bilingual group included bilinguals who learned English during 11-year period equivalent to early school years up until high school. With an additional group, the middle bilingual group could cover ages seven to twelve, and the late bilingual group could cover the period after puberty and on, 13 years and on.

Future research should consider a proficiency rating scale so each bilingual can indicate their exposure to both languages. Every individual reported when their AoA of English, but their current proficiency in both Spanish and English could be an additional factor that may affect their overall performance. Although all individuals were initially Spanish-speakers, many may not continue using the language due to an emphasis on English in school and their surrounding environment. This would affect their English proficiency, which may affect their performance in perception tasks based on English. Therefore, it would be useful to investigate if there is any correlation with language proficiency and the individual's performance.

Conclusion

The current study demonstrated that the age of acquisition and exposure to English affected the speech perception of bilingual speakers. Although there was not an English control to measure the overall effects of a second language, there were differences between the groups. The simultaneous and early bilingual groups performed similarly to each other. These similarities could be attributed to the similarities in language development for bilinguals who learn two languages from birth and up to six years of age. However, these similarities were not observed with the late bilinguals who had lower performance than the early bilinguals.

This study highlights the importance of being aware of the difficulties in speech perception of English for bilingual population. This will be important when considering the adverse listening conditions in work and school environments. Being aware of these difficulties will also be impactful when administering hearing assessments to Spanish-English bilingual individuals. The performance for these bilinguals may not be as accurate in English due to speech perception difficulties.

References

Birdsong, D. (1999). *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, N.J.: Routledge.

Birdsong, D. (2006). Age and second language acquisition and processing: a selective review. *Language Learning,* 56(s1), 9-49.

Bovo, R. & Callegari, E. (2009). Effects of classroom noise on the speech perception of bilingual children learning in their second language: Preliminary results. *Audiological Medicine*, 7(4), 226-232.

Brouwer, S., Van Engen, K. J., Calandruccio, L., Bradlow, A. R. (2011). Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. *J. Acoust. Soc. Am.*, 131(2), 1449-1464.

Broersma, M. & Scharenborg, O. (2010). Native and non-native listeners' perception of English consonants in different types of noise. *Speech Communication*, 52(11-12), 980-995.

Brungart, D. S. (2000). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.*, 109(3), 1101-1109.

Callahan, R. M., & Gándara, P. C. (Eds.). (2014). *Bilingual Education and Bilingualism, Volume 99: Bilingual Advantage: Language, Literacy and the US Labor Market*. Clevedon, GBR: Channel View Publications.

Cooke, M., Garcia Lecumberri, M. L., & Barker, J. (2007). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *J. Acoust. Soc. Am.,* 123(1), 414-427.

Cutler, A. Garcia Lecumberri, M. L. & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *J. Acoust. Soc. Am.*, 124(1264-1268), 1264-1268.

Garcia Lecumberri, M. L., Cooke, M., Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, 52(11–12), 864-886.

Gillam, R. B., Marquardt, T., Martin, F. N. (2011). *Communication sciences and disorders: From science to clinical practice*. Sudbury: Jones and Bartlett Publishers, LLC.

Guan, J. et al. (2015). Vowel identification in temporal-modulated noise for native and non-native listeners: Effect of language experience. *J. Acoust. Soc. Am.,* 138(3), 1670-1677.

Hisagi, M., Garrido-Nag, K., Datta, H., Shafer, V. L. (2015). ERP indices of vowel processing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 18(2), 271-289.

Jin, Su-Hyun & Liu, Chang (2014). English vowel identification in quiet and noise: effects of listeners' native language background. *Frontiers in Neuroscience*, 8(305), 1-9.

Liu, C. & Eddins, D. A. (2008). Categorical dependence of vowel detection in long-term speech-shaped noise. *J. Acoust. Soc. Am.,* 123(6), 4539-4546.

Lu-Feng, S. (2014). Speech audiometry and Spanish–English bilinguals: Challenges in clinical practice. *American Journal of Audiology*, 23(3), 243-258.

Marchegiani, L. & Fafoutis, X. (2015). One cross-language consonant identification in second language noise (L). *J. Acoust. Soc. Am.,* 138(4), 2206-2209.

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research,* 40, 686-693.

Meyer, A. & Schiller, N. (2003). *Phonetics and phonology in language comprehension and production, differences and similarities*. Berlin: Mouton de Gruyter.

Mi, L., et al. (2013). English vowel identification in long-term speech-shaped noise and multi-talker babble for English and Chinese listeners. *Acoustical Society of America*, 133(5), 391-397.

Morrill, T., Baese-Berk, M., Heffner, C., & Dilley, L. (2015). Interactions between distal speech rate, linguistic knowledge, and speech environment. *Psychon Bull Rev*, 22(5), 1451-1457.

Rhebergen, K. S., Versfeld, N. J., Dreschler, W. A. (2005). Release from informational masking by time reversal of native and non-native interfering speech (L). *J. Acoust. Soc. Am*. 118(3), 1274-1277.

Rosenhouse, J., Haik, L., & Kishon-Rabin, L. (2006). Speech perception in adverse listening conditions in Arabic-Hebrew bilinguals. *International Journal of Bilingualism*, 10(2), 119-135.

Shi, L. & Sanchez, D. (2010). Spanish/English bilingual listeners on clinical word recognition tests: What to expect and how to predict. *Journal of Speech, Language, and Hearing Research*, 53(5), 1096-1110.

Silbert, N. H., de Jong, K., Regier, K., Albin, A., Hao, Y. (2014). Acoustic properties of multi-talker babble. *J. Acoust. Soc. Am.*, 135(4), 2227.

Simpson, S. A. & Cooke, M. (2005). Consonant identification in N-talker babble is a nonmonotonic function of N (L). *J. Acoust. Sc. Am.,* 118(5), 2775-2778.

Snell, K. B. et al. (2002). Word recognition in competing babble and the effects of age, temporal processing, and absolute sensitivity. *J. Acoust. Soc. Am.,* 112(2), 720-727.

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J. Speech Hear. Res.,* 28, 455-462.

SykofizX Version 2.0 [Computer Software] (2003). Alachua, FL : Tucker-Davis Technologies.

Tabri, D., Smith Abou Charca, K. M., & Pring, T. (2011). Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language & Communication Disorders,* 46(4), 411-422.

Vineetha, C. V., Suma, R., Nair, S. P. (2013). Effect of bilingualism on speech in noise perception in young adults. *Language in India*, 13(6), 799-811.

Watson, C. S. (2005). Some comments on informational masking *Acta Acustica United with Acustica*, 91, 502-512.

Wilson, S. M., Iacoboni, M. (2006). Neural responses to non-native phonemes varying in productability: Evidence for the sensorimotor nature of speech perception. *NeuroImage*, 33(1), 316-325.